ISYE 320 Midterm Report

April 4th, 2024

Eric Stratman

By: Ty Mountain, Nandini Jagtiani, Aditya Narasimhan, Olivia Balara

Table of Contents

1. Introduction and Problem Statement

1.1 The System

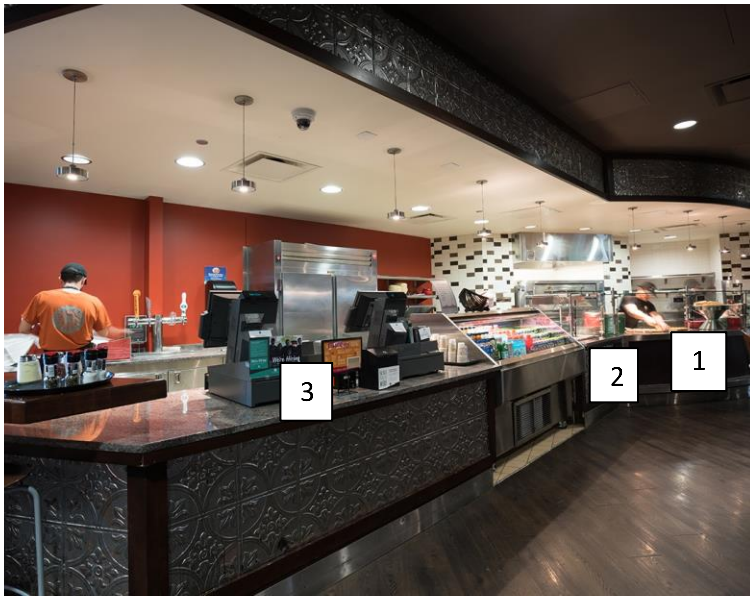

Figure 1: Image of System with Labeled Stations

The team chose to model South Cantina, a university owned restaurant in the Union South building on the University of Wisconsin-Madison campus. South Cantina is a counter style restaurant with three primary sequential stations that processes food orders. The customer follows the food being prepared from station to station. The queue and following processes are all easily observed, making data collection simple and accurate. The three stations include a protein/base station (1), a toppings station (2), and a register station (3) as seen in *Figure 1*. A fourth station is occasionally added between the toppings and register station, where an additional employee assists in the closing of a burrito bowl or the wrapping of a tortilla. Each station is manned by a single employee. This process is fed through the base station by a line of customers waiting in a queue. The team will study the system at lunchtime from 11 am to 2 pm on weekdays.

Customers leaving the queue begin by telling the employee at the base station whether they would like a burrito, a burrito bowl, taco, nachos, or salad. The employee picks up the respective container and fills it with the customer's base selection. This employee then hands the container to their right to the next employee manning the toppings station. During times of high demand, WIP is held on counters between employees.

At the toppings station, the employee asks the customer what toppings they want. The customer has an option to choose from both standard toppings and premium toppings. With all orders, customers can choose an unlimited number of standard toppings. The employee will add the selected toppings to the container and either close the container themselves, or hand it to their right to the next employee manning the occasional closing station. Either the topping station employee or the closing station employee then closes the container with the respective lid or wraps the burrito with foil, which is then handed to the customer.

Once the customer has received their order, they move to the register. The employee at the register station asks the customer if they selected any premium toppings and asks them if they would like a drink. After calculating the total amount, the employee asks the customer what method they would like to pay with and offers to give the receipt after the payment is approved. The customer can then exit the restaurant.

1.2 Current Challenges

South Cantina, a bustling counter-style restaurant, encounters several operational challenges that affect service efficiency, particularly during peak lunch hours from 11 am to 2 pm during weekdays. The observed challenges include:

1. Variable wait times: Customers experience fluctuating wait times, especially during the period breaks aligned with class schedules, leading to sudden surges in the queue.

2. Service time inconsistencies: The time taken to serve customers varies significantly depending on the type of order. Burritos require a longer preparation time compared to bowls, nachos, and tacos, contributing to an uneven flow in the service process.

3. Staffing flexibility: Occasionally, the arrival of an employee at the closing station allows the toppings station employee to process orders faster, as they no longer must close/wrap the item. This leads to inconsistent service times.

4. Queue management: During peak times, the queue can become unruly and difficult to manage, which can lead to customer dissatisfaction and a potential decline in service quality.

1.3 Why Simulation is the Right Tool to Address These Challenges

Simulation presents an optimal tool for addressing these challenges due to its ability to model complex systems and test changes without disrupting actual operations. Through simulation, we can:

1. Analyze the process flow to identify bottlenecks and understand the impact of variable demand on wait times.

2. Experiment with different staffing arrangements to determine how best to accommodate varying preparation times for different menu items.

3. Examine the effects of the “Closing Station” and how its inclusion affects key performance indicators.

4. Test queue management strategies to enhance customer flow and reduce wait times.

2. Conceptual Model and Assumptions

2.1 The Conceptual Model

The team will model South Cantina by representing how customer arrivals flow through the three primary processes: the base station, toppings station, and closing station. The occasional use of the closing station will not be included in this simulation as not enough data was taken to determine the utilization of the station and its resources. The addition of this station is often erratic, with a fourth employee aiding in the finishing of food items during times of peak demand. The addition of this station will be a primary focus of future experiments with the simulation.

The below figure describes the conceptual model of the system.

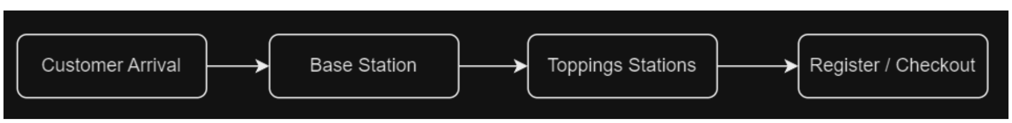

Figure 2: Conceptual Model of South Cantina

The team will use the Arena software package to create a computer model of the system shown in *Figure 2*. In Arena, customer arrivals will be simulated using a create module. This module will create entities according to a determined distribution that can flow through the remaining processes. The following stations: base, toppings, and register, will each be represented by a process module. As previously mentioned, each of these stations has a dedicated employee who can work on one order at a time. This will be reflected in the simulation by giving each station its own resource with a capacity of one, showing that each employee can only process one customer at a time. Each process module will operate as a “Seize Delay Release”. This will allow the model to identify when the employees (resources) are performing a task. The seize function will halt all incoming customers from progressing to the station since the resource is currently in use and will release the resource upon task completion which is represented by the delay.

Following their arrival, the customers (entities) will move on to the base station to pick what type of meal they will be ordering, i.e. burrito, burrito bowl, etc. and adding the protein of their choice. After this selection, they move on to the toppings station and pick their desired toppings.

The final station in the model is the register station. A dispose module will follow the register module to simulate customers leaving the system.

2.2 Assumptions

- Queue Discipline:
    - We assumed that the customers are served on a FIFO (first in first out) basis. Since the process is done in serial, the first person that enters the system is typically the first person to leave the system.
- Queue Behavior:
    - There were no instances of balking, reneging, or jockeying. This is because there was only one queue. There were also no instances of customers leaving in the middle of the queue.
- System Capacity:
    - There is no capacity. Hypothetically there could be an infinite number of customers in the system.
- Calling Population:
    - Calling population is infinite. The number of entities that can enter a system is infinite or significantly large.
- Continuous Data:
    - While data points are integers, the data collected is still categorized as continuous since time is being measured.

2.3 Sources of Randomness

The simulation shows many sources of randomness. Some of these include:

- Customer Interarrival Time and Group Size:
    - While we are examining data within a busier time frame with higher demand, the rate at which customers enter the system is random. It is also not guaranteed that customers will enter the system alone.
- Customer Order Type:
    - South Cantina offers 5 different menu items, as mentioned before, each that takes a different amount of time to complete. For example, the burrito process has a longer completion time at the base station since it involves heating up the tortilla before adding the protein and toppings.
- Topping Selection and Amount:
    - Each customer picks a unique combination and a variable number of toppings to be added to their base. This process thereby sees a lot of variation in its completion time since customers can add anywhere from 1-2 toppings to all the toppings available to their base.
- Premium Toppings:

    - While most toppings offered are available for all menu items, the queso and guacamole topping incur an extra charge. When choosing a premium topping the employee mentions that they cost extra and reflect it upon the item's packaging so that the customer is charged appropriately at the register station.
- Employee Proficiency:
    - Employee proficiency is also a source of randomness applicable to the model. Variation in experience between employees can affect the speed at which all stations operate. Experienced employees might be more familiar with the process and therefore run the process faster, whereas less experienced employees might not be able to operate at the same speed.
- Food Replacement:
    - Another source of randomness in the model is the food replacement procedure. The protein and toppings are placed in large metal containers which are replaced when they run out. Often this happens in the middle of customer service, causing variation in task completion time.

3. Data Collection and Input Analysis

3.1 Data Collection Methods and Conditions

Data was collected between 11:10 am and 2:00 pm over 4 different days. Data collection was performed during weekdays, specifically once on Wednesday, twice on Thursday, and once on Friday.

The first day of data collection consisted of interarrival times over 11am to 1pm. During this period, all arrivals were taken to control for potential non-homogenous arrival rates that may result from period changes. This process yielded 102 arrival time data points. Groups were counted as a single arrival. This was only relevant for a few datapoints, primarily in cases where groups of young children entered the queue.

This data collection was performed by marking time stamps from a running clock. The same method was used for the other three days, however interarrival times were only taken on day 1. Data for the closing station (Labelled under “Employee 3?”) was taken if that station was present, however it was not used during model construction (*Fig. 3*).

			Employee 1		Employee 2		Employee 3?		Employee 4				
#	Description	Arrival Time	Protein/Base Start	Protein/Base End	Toppings Start	Toppings End	Container Close Start	Container Close End	Register Start	Register End	Contiguous?	Bad Data?	Comments
	Enter info about customer	*Arrive to queue area.*	*1st employee talks to customer*	*Food moves to topping area*	*2nd employee begins talking OR first topping grabbed*	*Food handed to customer or handed to third employee*	*A third employee begins putting lid on container*	*Food handed to customer*	*customer begins talking to casher*	*customer walks away*	*Indicate if after pervious customer on data sheet*	*Make note if should ignore?*	...
1		1:30	1:30	2:00	2:00	2:43			2:50	3:03			
2		7:13	7:16	7:55	7:55	8:40			8:45	9:20	x		
3		11:18	11:18	12:09	12:10	12:29			12:34	12:55	x		
4		20:13	20:13	20:38	20:38	21:27			21:31	22:00	x		
5		20:13									x		
6		20:55									x		

Figure 3: Sample of Data Collection Sheet from Day 1

A sample of the final data collection table is shown in *Figure 4*. 102 data points were collected for the interarrival times and 50 data points were collected for each of the base, toppings, and register stations.

No.	Interarrival Times	Base Station	Toppings Station Only	Register Station
1	90	30	43	13
2	343	39	45	35
3	245	51	19	21
4	535	25	49	29
5	0	37	86	24
6	42	29	34	29
7	0	30	37	28
8	100	14	29	20
9	70	23	18	26
10	15	21	14	19
11	12	37	83	18
12	56	36	55	12
13	237	20	52	15
14	184	7	27	7
15	86	15	40	21
16	27	16	14	20
17	160	16	59	26
18	0	27	33	16
19	112	28	34	15
20	58	17	38	9
21	57	12	20	15
22	50	23	50	19
23	4	26	67	6
24	44	18	42	8
25	11	37	68	27
26	53	57	34	15
27	1	42	27	10
28	10	17	67	24
29	0	41	40	24
30	23	23	30	23
31	22	39	24	23

Figure 4: Sample of Final Data Points in Seconds for Each Process

3.2 Data Analysis

3.2.1 Arrivals

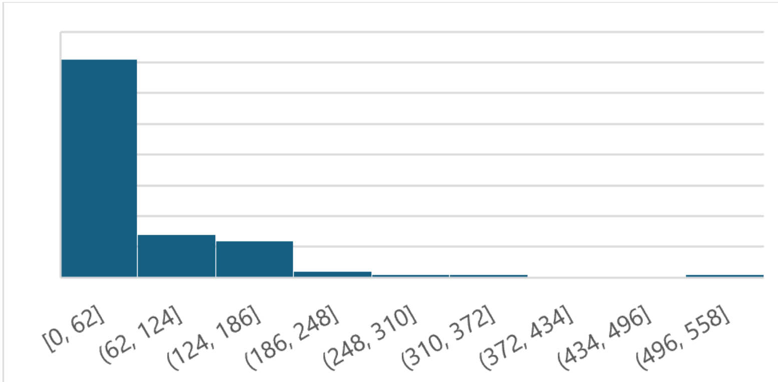

Figure 5: Histogram of the Interarrival Times

The histogram of interarrival times in *Figure 5* illustrates the frequency with which customers arrive at South Cantina. The majority of interarrival times fall within the shortest time bin, indicating that customers often arrive in quick succession. As the interarrival time increases, the frequency of occurrences significantly drops, suggesting that longer waits between arrivals are less common.

Given that the interarrival times at South Cantina are recorded in seconds, the histogram indicates that the data could follow an exponential distribution, which is characterized by its memoryless property. This implies that the probability of a customer's arrival is not affected by the time elapsed from the previous arrival. The observed pattern, where a higher frequency of customers arrive after shorter intervals, supports the use of an exponential model. The distribution's shape indicates that most customers arrive after a short period following the previous customer, which is typical for busy eateries during peak hours.

Summary Statistics	
Mean	62.46078
Median	31.5
Standard Deviation	83.46056

Figure 6: Summary Statistics of the Interarrival Times

The interarrival times at South Cantina, with an average of approximately 62.5 seconds and a median of 31.5 seconds, indicate a right-skewed distribution. This skew suggests customers frequently arrive in clusters, with a significant number occurring in shorter intervals. The standard deviation of 83.46 seconds points to a wide variation in the time between arrivals, highlighting the irregular flow of customer traffic during the non-peak hours seen in *Figure 6*.

```
        Distribution Summary

Distribution:   Exponential
Expression:     -0.001 + EXPO(62.5)
Square Error:   0.022473

Chi Square Test
  Number of intervals   = 3
  Degrees of freedom    = 1
  Test Statistic        = 10.4
  Corresponding p-value < 0.005

Kolmogorov-Smirnov Test
  Test Statistic        = 0.153
  Corresponding p-value = 0.0166

        Data Summary

Number of Data Points   = 102
Min Data Value          = 0
Max Data Value          = 535
Sample Mean             = 62.5
Sample Std Dev          = 83.3

        Histogram Summary

Histogram Range         = -0.001 to 535
Number of Intervals     = 10
```

Figure 7: Data Distribution Summary for the Interarrival Times

Goodness of Fit Test (Chi-Squared Test)

We use the Chi Squared Test in this case, since we have more than 20 data points.

Hypothesis:
H_o: The chosen distribution fits the data (null hypothesis)
H_1: The chosen distribution does not fit the data (alternative)

Test statistic X_0^2=10.4, (*Fig. 7*)
Using a 95% Confidence Interval, $\alpha = 0.05$, df = 1
Critical Value $X_{\alpha,k-s-1}^2$= 3.84

Result: Since our $X_0^2 > X_{\alpha,k-s-1}^2$, we reject our null hypothesis H_o and can conclude that the data does not theoretically follow an exponential distribution.

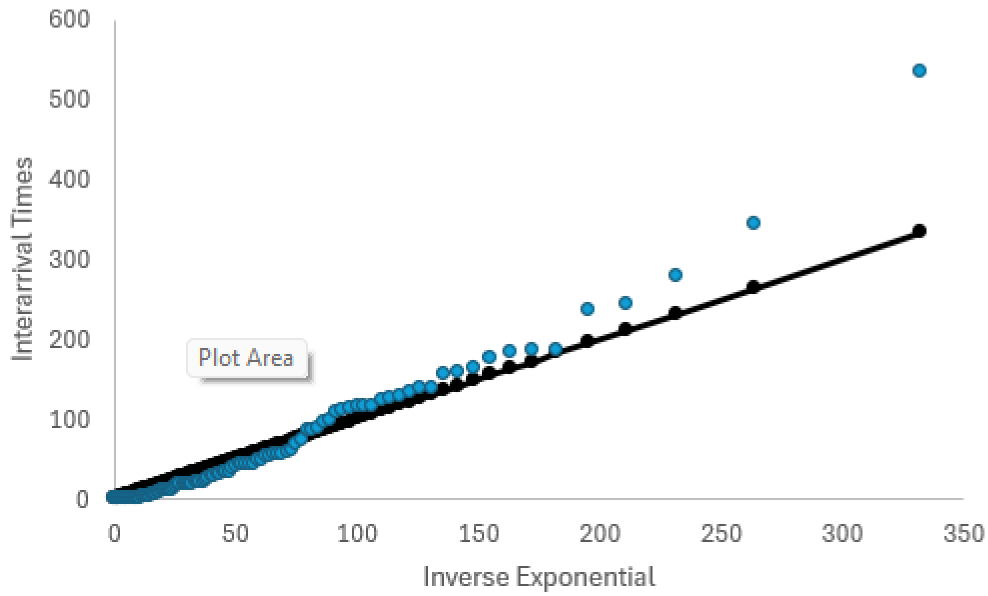

Figure 8: Interarrival Times QQ Plot

Looking at the QQ plot in *Figure 8*, most data points lie close to the 45-degree reference line, with only a few outliers deviating from this pattern. This suggests that the data may reasonably follow an exponential distribution. Collecting a larger sample size could help confirm whether this distributional assumption holds more definitively. The result of the QQ plot provides enough evidence to use the exponential distribution for arrivals despite the failed chi-squared test.

3.2.2 Base Station

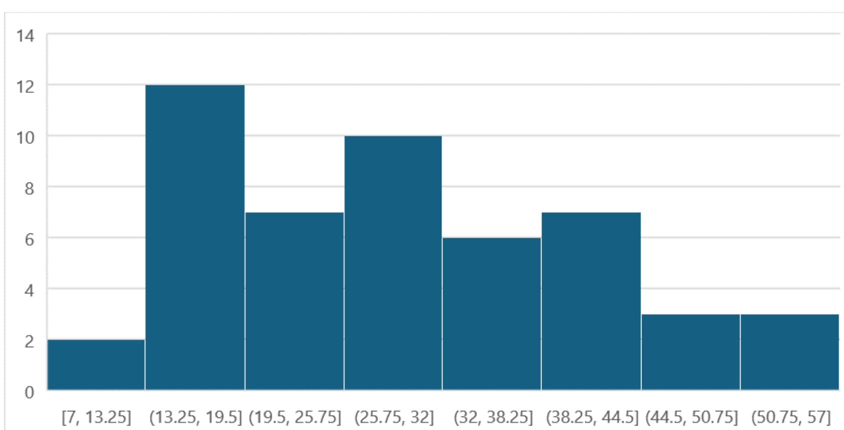

Figure 9: Histogram of the Base Station Time Distribution

The base station time histogram in *Figure 9* shows that most customer interactions with the employee at the base station are brief, falling within the [13.25,19.5] interval. Frequency decreases as base station times increase, indicating that longer interactions are less common. However, a slight uptick in the [38.25, 44.5] interval suggests that some customer orders, such as burritos, may require longer preparation times compared to other menu items. The right-skewed distribution implies that while most service times are short, there is a tail of longer durations.

The shape of the base station time histogram suggests that the data might follow a normal distribution. The graph shows a concentration of data points around the central intervals, with the frequency decreasing as we move towards the tails on either side. This symmetric, bell-shaped pattern is characteristic of a normal distribution.

Sample Mean	29.26
Median	27.5
Sample Standard Deviation	12.08272506

Figure 10: Summary Statistics of Base Station Times

The average time spent at the base station is approximately 29.3 seconds, with a median nearly equal to the mean, indicating a negligible skew in the distribution. A standard deviation of 12.1 seconds signifies a modest spread in times, which suggests a consistent duration of base station visits among customers in *Figure 10*. The negligible skew also hints at normal distribution.

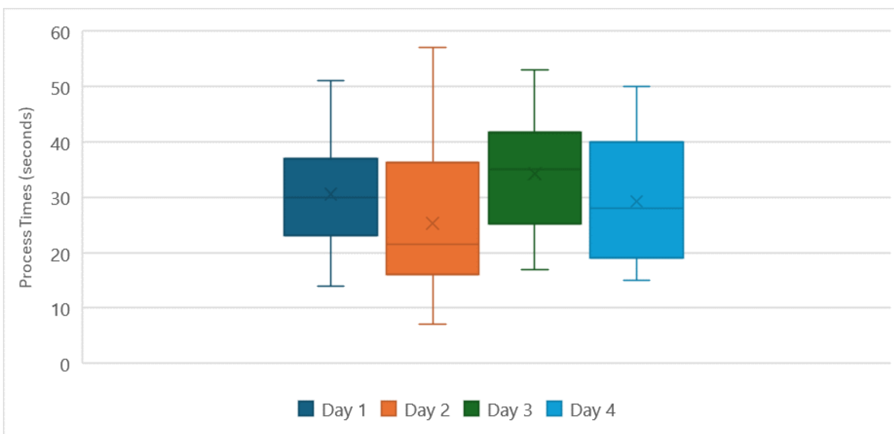

Figure 11: Box Plot for Each Day of Base Station Data Collection

The box plot in *Figure 11* allows the team to visually assess the homogeneity of the base station time data across the four days. While there are some differences in the medians and spreads, the overall distributions seem similar. The interquartile ranges (IQRs) of the four days overlap considerably, and there are no extreme outliers present. The relatively few data points from each day make it more likely to see slight differences between each box plot.

```
        Distribution Summary

Distribution:   Normal
Expression:     NORM(29.3, 12)
Square Error:   0.020309

Chi Square Test
  Number of intervals   = 8
  Degrees of freedom    = 5
  Test Statistic        = 4.82
  Corresponding p-value = 0.449

        Data Summary

Number of Data Points   = 50
Min Data Value          = 7
Max Data Value          = 57
Sample Mean             = 29.3
Sample Std Dev          = 12.1

        Histogram Summary

Histogram Range         = 6.5 to 57.5
Number of Intervals     = 51
```

Figure 12: Distribution Summary of the Base Station Times

Goodness of Fit Test (Chi-Squared Test)
Hypothesis:
H_o: The chosen distribution fits the data (null hypothesis)
H_1: The chosen distribution does not fit the data (alternative)

Test statistic X_0^2= 4.82, (*Fig. 12*)
Using a 95% Confidence Interval, $\alpha = 0.05$, df = 5
Critical Value $X_{\alpha,k-s-1}^2$= 11.1

Result: Since our $X_0^2 < X_{\alpha,k-s-1}^2$, we do not reject our null hypothesis H_o and can conclude that the data follows a normal distribution.

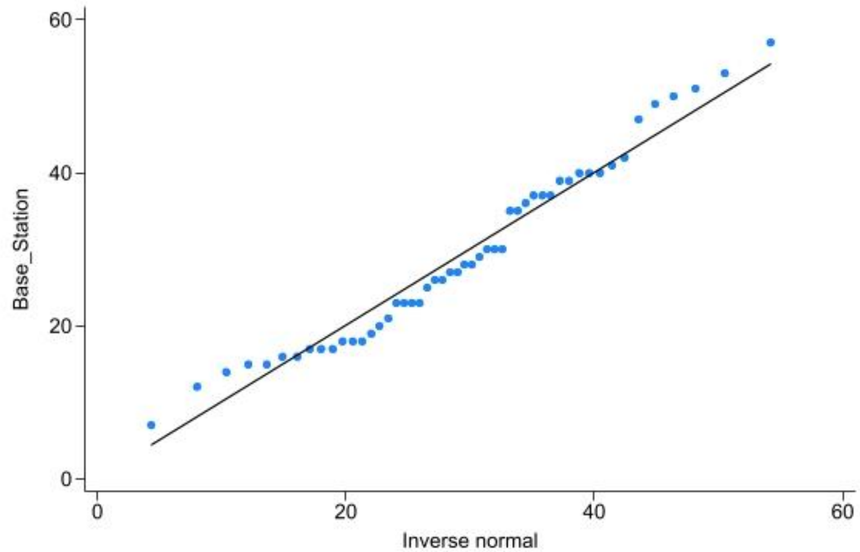

Figure 13: Base Station QQ Plot

The QQ plot in *Figure 13* displays a linear pattern with the data points closely following the 45-degree reference line, which suggests that the base station times are approximately normally distributed. There are no significant deviations from the line, reinforcing the conclusion of little to no skew in the data.

3.2.3 Toppings Station

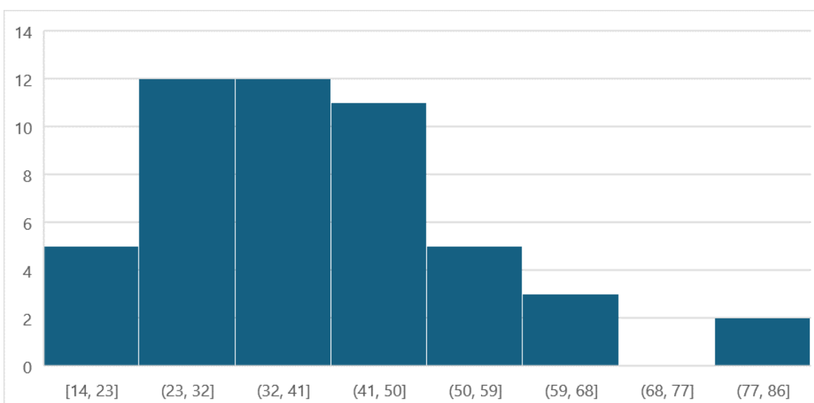

Figure 14: Toppings Station Time Distribution

The toppings station histogram, *Figure 14* reveals a distribution of customer service times with a clear concentration in the middle intervals, particularly within the [23, 41] second range. This suggests that the average service time at the toppings station is moderately long compared to the base station, with these intervals capturing the peak of the distribution. The histogram displays a gradual decrease in frequency as we move towards the higher and lower time intervals, demonstrating that extremely short or long topping times are less common.

The shape of the toppings station histogram, with most times grouped in the middle, suggests that the service times could be normally distributed despite a small bunching near the mean. This regular pattern indicates that the duration of customer visits at the toppings station is usually predictable and consistent during lunch time.

Sample Mean	40.2
Median	38.5
Sample Standard Deviation	16.0814255

Figure 15: Summary Statistics of Toppings Station Times

As shown in *Figure 15*, The mean time at the toppings station is 40.2 seconds with a very close median of 38.5 seconds, which implies that the data is nearly symmetrically distributed with negligible skew. The closeness of the mean and median is a strong indicator of symmetry in the distribution, as a significant skew would have pulled the mean away from the median. The standard deviation of 16.08 seconds indicates variability in the topping station times, yet this variability is consistent with a normal distribution, given the minimal skew reflected by the proximity of the mean to the median.

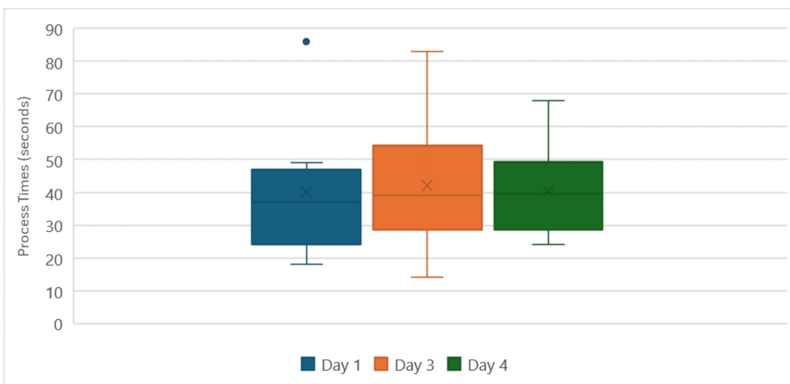

Figure 16: Box Plot for Each Day of Toppings Station Data Collection

The box plot in *Figure 16* shows service times at the toppings station across three distinct days, with the data appearing symmetrical around the median denoted by the lines within each box. This symmetry is underscored by the proximity of the 'X' marks, representing the means, to the median lines. The uniformity of the interquartile ranges (IQRs), which are the heights of the boxes, indicates a consistent distribution of process times across the days.

Notably, there is an outlier on Day 1; however, this single data point does not seem to disrupt the overall symmetry of the distribution. The alignment of the medians and means, coupled with the symmetrical nature of the boxes, lends credence to the claim of a normal distribution. This consistency and symmetry are indicative of homogeneous data, as would be expected in a stable and predictable process. The box plot, therefore, provides a visual confirmation that the toppings station's service times are not only normally distributed but also exhibit homogeneity across the different days.

```
        Distribution Summary

Distribution:   Normal
Expression:     NORM(40.2, 15.9)
Square Error:   0.017324

Chi Square Test
  Number of intervals   = 8
  Degrees of freedom    = 5
  Test Statistic        = 8.5
  Corresponding p-value = 0.142

        Data Summary

Number of Data Points   = 50
Min Data Value          = 14
Max Data Value          = 86
Sample Mean             = 40.2
Sample Std Dev          = 16.1

        Histogram Summary

Histogram Range         = 13.5 to 86.5
Number of Intervals     = 73
```

Figure 17: Distribution Summary of Toppings Station Times

Goodness of Fit Test (Chi-Squared Test)

Hypothesis:
H_o: The chosen distribution fits the data (null hypothesis)
H_1: The chosen distribution does not fit the data (alternative)

Test statistic X_0^2= 8.5, (*Fig. 17*)
Using a 95% Confidence Interval, $\alpha = 0.05$, df = 5
Critical Value $X^2_{\alpha,k-s-1}$= 11.1

Result: Since our $X_0^2 < X^2_{\alpha,k-s-1}$, we do not reject our null hypothesis H_o and can conclude that the data follows a normal distribution.

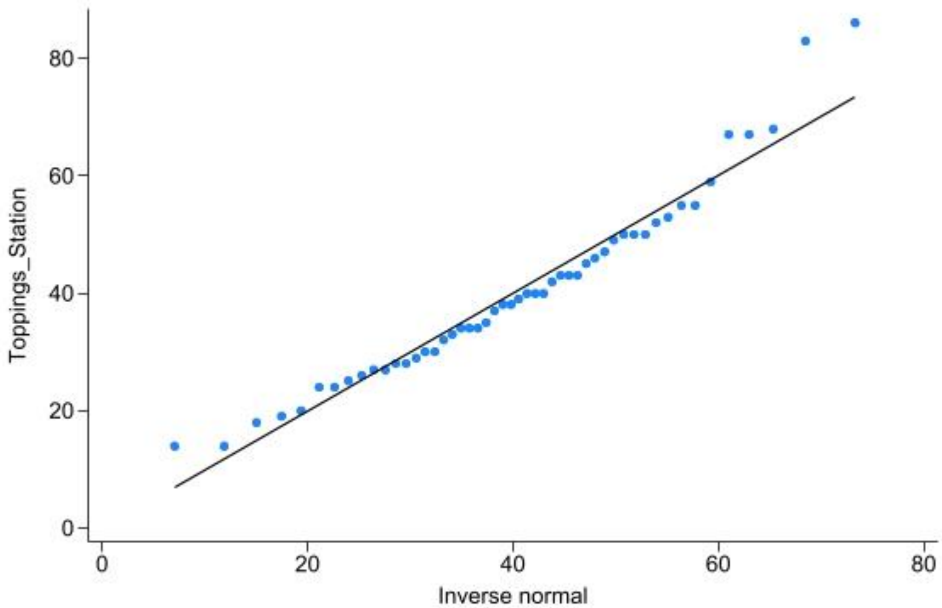

Figure 18: Toppings Station QQ Plot

The QQ plot in *Figure 18* for the Toppings Station shows that most data points line up along the reference line, indicating that the service times are mostly consistent with a normal distribution. There are slight deviations at the ends, but these do not heavily impact the overall normality of the data.

3.2.4: Register Station

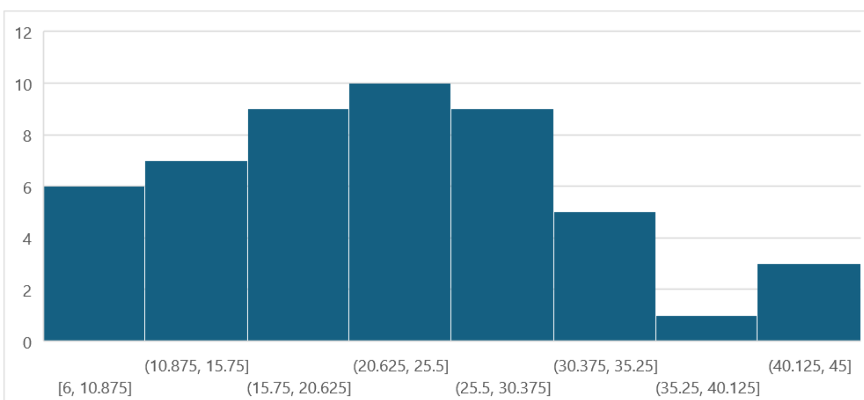

Figure 19: Register Station Time Distribution

The histogram of the register station (*Fig. 19*) time distribution shows a concentration of service times within the 15.75 to 30.375-second range, with the peak frequency in the 20.625 to 25.5-second interval. The shape of the histogram is somewhat symmetrical around this peak interval, which could indicate that the service times at the register station are normally distributed. Despite a slight decrease in frequency at the tails, particularly for longer times, the distribution does not show a pronounced skew. This balance in the spread of times on both sides of the most common interval supports the assertion of a normal distribution pattern.

Sample Mean	22.66
Median	23
Sample Standard Deviation	9.403885432

Figure 20: Register Station Summary Statistics

The mean service time at the register station is 22.66 seconds, with a median very close at 23 seconds, suggesting a distribution with minimal skew. The proximity of the mean to the median in *Figure 20* indicates that the service times are symmetrically distributed around the central value, which is a characteristic of a normal distribution. A standard deviation of 9.4 seconds shows there's some variability in the service times, but this amount of spread is not unusual for a normal distribution.

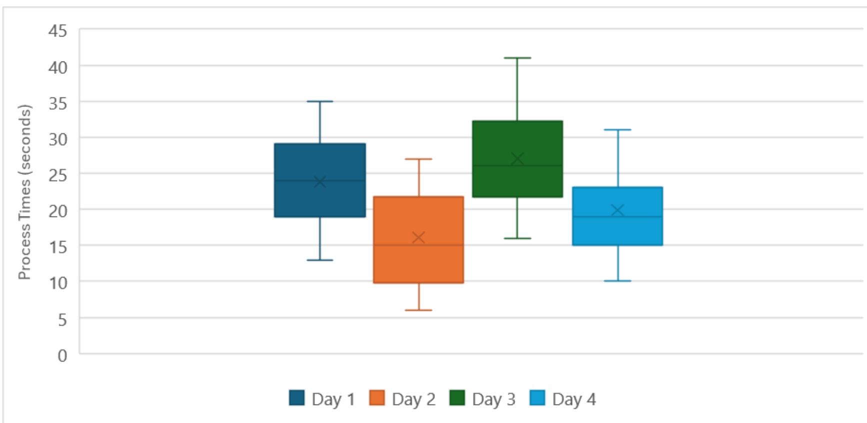

Figure 21: Box Plot for Each Day of Register Station Data Collection

The box plot (*Fig. 21*) displays some variability in the register station times across the four days, but the overall distributions exhibit a fair degree of consistency. The medians range from about 15 seconds (Day 3) to around 30 seconds (Day 1), and the interquartile ranges of the four days overlap. Additionally, there are no extreme outliers present on any of the days. Despite the differences in medians and ranges, the box plot suggests a reasonable level of consistency in the data across the four days.

```
        Distribution Summary

Distribution:   Normal
Expression:     NORM(22.7, 9.31)
Square Error:   0.018353

Chi Square Test
  Number of intervals   = 8
  Degrees of freedom    = 5
  Test Statistic        = 3.13
  Corresponding p-value = 0.683

        Data Summary

Number of Data Points   = 50
Min Data Value          = 6
Max Data Value          = 45
Sample Mean             = 22.7
Sample Std Dev          = 9.4

        Histogram Summary

Histogram Range         = 5.5 to 45.5
Number of Intervals     = 40
```

Figure 22: Distribution Summary of the Register Station

Goodness of Fit Test (Chi-Squared Test)

Hypothesis:
H_o: The chosen distribution fits the data (null hypothesis)
H_1: The chosen distribution does not fit the data (alternative)

Test statistic X_0^2= 3.13 (*Fig. 22*)
Using a 95% Confidence Interval, $\alpha = 0.05$, df = 5
Critical Value $X_{\alpha,k-s-1}^2$= 11.1

Result: Since our $X_0^2 < X_{\alpha,k-s-1}^2$, we do not reject our null hypothesis H_o and can conclude that the data follows a normal distribution.

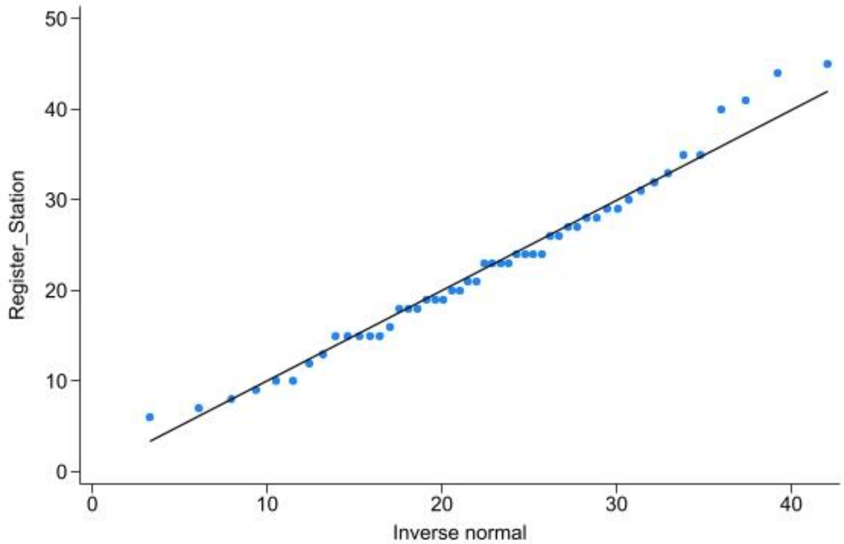

Figure 23: Register Station QQ Plot

Figure 23 shows a strong linear relationship between the observed data and the expected normal values for most of the range, indicating that the data closely follows a normal distribution. However, there is a slight deviation from the linear trend at both the lower and upper extremes, suggesting potential departures from normality in the tails of the distribution. Overall, the QQ plot supports the assumption of approximate normality for most of the data.

3.2.5 Independence

As a final step of data analysis, the team calculated Pearson coefficients to determine whether any correlation existed between processes. These values are displayed in *Figure 24*, with values near 1 or -1 indicating correlation while values near zero indicate independence. All values below are near enough to zero that the team concluded that correlation was not present between processes.

	Interarrivals	Base Station	Toppings Station	Register Station
Interarrivals		-0.08459498	-0.020558437	-0.010316356
Base Station			0.023253702	0.218519148
Toppings Station				0.015757826

Figure 24: Pearson Coefficient Table Between Processes

4. Model Translation, Verification, and Validation

4.1 Model Translation

The conceptual model was translated into a computer simulation using the Arena software package. The entities that flow through the model represent customers at South Cantina. These entities enter the system through a create module with a delay between arrivals of (EXPO(62.5)-.001), which is the formula given by the input analyzer (*Fig. 7*). Units for this module and the following process modules were all set to seconds.

Three process modules are used to represent the base, toppings, and register station. All three of these processes are arranged in series. Each process has its own resource, each with a capacity of one. The delay for each process was set to normal with the mean and standard deviation set to the values from the input analyzer as summarized in *Figure 25*. Each process operates as a "Seize Delay Release" which allows each entity to seize the resource available at each station until the delay is complete.

Station	Mean	Standard Deviation
Base	29.3	12
Toppings	40.2	15.9
Register	22.7	9.31

Figure 25: Normal Distribution Parameters of the Process Modules

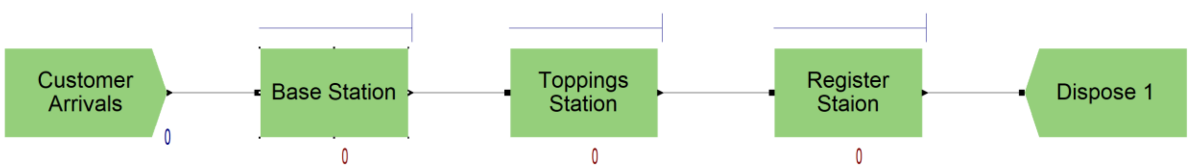

Figure 26: Arena Translation of the Conceptual Model

The model (*Fig. 26*) aims to simulate South Cantina over the period the team observed it. Data was taken between 11:10am and 2:00pm on weekdays, therefore the model will use a warm-up period of 10 minutes and have a replication length of one three-hour day. Because we studied the system on weekdays, we will simulate one week by using 5 replications when analyzing data on the system. Upon running these replications, entities move through the model, and it is visually stable as queues are not perpetually growing at any of the processes. Validation through comparison with new output data is required to confirm whether the simulation accurately simulates real world conditions.

4.2 Model Validation

Statistical model validation was performed by comparing the average number of customers in the system with 17 data points collected after the simulation was created and verified. These data points were taken from 12:40 pm to 2:00 pm in intervals of 5 minutes. They are shown below in *Figure 27*.

Number of customers	Time
7	12:40
5	12:45
4	12:50
5	12:55
3	1:00
3	1:05
1	1:10
4	1:15
5	1:20
3	1:25
1	1:30
2	1:35
1	1:40
1	1:45
0	1:50
0	1:55
1	2:00

Figure 27: Validation Data of Customers in System

To validate the simulation, the team utilized a replication-deletion method of creating data points to average and compare with the validation data. 30 replications (days) of the simulation were run and the average number of entities in the system was averaged across these replications. A test was performed with the following hypotheses:

H_o: The validation and simulation means are equivalent (null hypothesis)
H_1: The validation and simulation means are not equivalent (alternative)

This test is performed to determine whether the difference between these means could be generated by a normal random variable with a mean of zero. This was conducted by performing a t-test with an alpha value of .05 (*Fig. 28*).

Validation Mean	2.705882353
Simulation Mean	2.201879683
n	30
Simulation Standard Deviation	0.25543865
Test Statistic	-10.80704235
Degrees of Freedom	29
Critical Value	2.04523

Figure 28: Validation T-Test Values

After comparing the absolute value of the test statistic with the critical value at an alpha of .05, the null hypothesis must be rejected as the test statistic is greater than the critical value.

The team wanted to determine how much the large number of replications contributed to this test's failure. Therefore, a second test was performed with only 15 replications to determine how close the null hypothesis was to holding (*Fig. 29*).

Validation Mean	2.705882353
Simulation Mean	2.224064441
n	15
Simulation Standard Deviation	0.24794082
Test Statistic	-7.526282866
Degrees of Freedom	14
Critical Value	2.14479

Figure 29: Second Validation T-Test Values

This second test yielded the same result, as the absolute value of the test statistic is still larger than the critical value. This confirms that our initial test is sound and not sensitive to the number of replications. Even at an extremely sensitive alpha value of .001, the test statistic still exceeds the critical value of 4.1405.

The team hypothesizes that the simulations' inability to be validated against external output data is largely the result of too few constructive data points as well as too few validation data points. Given the simplicity of the system and the small period that it was studied over, several hundred data points of both varieties would likely result in an accurate simulation. The system's construction and its process time distributions are suspected to be correct, but the accuracy of their parameters and the volume of validation data are areas of future improvement.

5. Next Steps

One system change that the team will experiment with is examining the effect of keeping the closing station open full time. As mentioned previously, South Cantina occasionally utilizes an extra employee to assist the toppings station employee with the preparation of the food item in a process known as the closing station. The team suspects that the addition of this station full-time will increase throughput as the cycle time of the toppings station will decrease on average as it is split between two employees. However, the effectiveness of this addition and whether any benefit is worth the cost of the extra employee needs to be studied. An element of this experiment that the team plans to analyze is how the addition of a closing station reduces the maximum queue length. This experiment will require collecting more data on the toppings and closing stations when the closing station is in use. A process module will need to be added to the simulation to represent the closing station, and the parameters of the toppings station will need to be adjusted based on the data collected when a closing station is in use.

A variation of this experiment that the team plans to study is how the benefits of the closing station change if it does not require its own employee but is performed by the same employee as the register station. In this scenario, the process may be able to realize some of the benefits of a closing station without incurring the cost of a fourth employee. Because the register station has the smallest mean process time of the three processes, the register employee can potentially stand to increase its utilization without becoming too severe of a bottleneck. This is effectively a method of line balancing. To perform this experiment, the closing station and the register station would need to share the same resource in the Arena simulation. The results of both experiments will be determined by examining key performance metrics such as the average number of customers in process, but more KPI's such as utilization and maximum queue length will be considered as well.